

# On Randomized Algorithms for Matching in the Online Preemptive Model

Ashish Chiplunkar                      Sumedh Tirodkar  
 {ashish.chiplunkar@gmail.com}      {sumedht@cse.iitb.ac.in}

Sundar Vishwanathan  
 {sundar@cse.iitb.ac.in}

## Abstract

We investigate the power of randomized algorithms for the maximum cardinality matching (MCM) and the maximum weight matching (MWM) problems in the online preemptive model. In this model, the edges of a graph are revealed one by one and the algorithm is required to always maintain a valid matching. On seeing an edge, the algorithm has to either accept or reject the edge. If accepted, then the adjacent edges are discarded. The complexity of the problem is settled for deterministic algorithms [6, 8].

Almost nothing is known for randomized algorithms. A lower bound of 1.693 is known for MCM with a trivial upper bound of two. An upper bound of 5.356 is known for MWM. We initiate a systematic study of the same in this paper with an aim to isolate and understand the difficulty. We begin with a primal-dual analysis of the deterministic algorithm due to [6]. All deterministic lower bounds are on instances which are trees at every step. For this class of (unweighted) graphs we present a randomized algorithm which is  $\frac{28}{15}$ -competitive. The analysis is a considerable extension of the (simple) primal-dual analysis for the deterministic case. The key new technique is that the distribution of primal charge to dual variables depends on the “neighborhood” and needs to be done after having seen the entire input. The assignment is asymmetric: in that edges may assign different charges to the two end-points. Also the proof depends on a non-trivial structural statement on the performance of the algorithm on the input tree.

The other main result of this paper is an extension of the deterministic lower bound of Varadaraja [8] to a natural class of randomized algorithms which decide whether to accept a new edge or not using *independent* random choices. This indicates that randomized algorithms will have to use *dependent* coin tosses to succeed. Indeed, the few known randomized algorithms, even in very restricted models follow this.

We also present the best possible  $\frac{4}{3}$ -competitive randomized algorithm for MCM on paths.

## 1 Introduction

Matching has been a central problem in combinatorial optimization. Indeed, algorithm design in various models of computations, sequential, parallel, stream-

ing, etc., have been influenced by techniques used for matching. We study the maximum cardinality matching (MCM) and the maximum weight matching (MWM) problems in the online preemptive model. In this model, edges  $e_1, \dots, e_m$  of a graph, possibly weighted, are presented one by one. An algorithm is required to output a matching  $M_i$  after the arrival of each edge  $e_i$ . This model constrains an algorithm to accept/reject an edge as soon as it is revealed. If accepted, the adjacent edges, if any, have to be discarded from  $M_i$ .

An algorithm is said to have a *competitive ratio*  $\alpha$  if the cost of the matching maintained by the algorithm is at least  $\frac{1}{\alpha}$  times the cost of the offline optimum over all inputs. The deterministic complexity of this problem is settled. For maximum cardinality matching (MCM), it is an easy exercise to prove a tight bound of two.

The weighted version (MWM) is more difficult. Improving an earlier result of Feigenbaum et al, McGregor [6] gave a deterministic algorithm together with an ingenious analysis to get a competitive ratio of  $3 + 2\sqrt{2} \approx 5.828$ . Later, this was proved to be optimal by Varadaraja [8].

Very little is known on the power of randomness for this problem. Recently, Epstein et al. [3] proved a lower bound of  $1 + \ln 2 \approx 1.693$  on the competitive ratio of randomized algorithms for MCM. This is the best lower bound known even for MWM. Epstein et al. [3] also give a 5.356-competitive randomized algorithm for MWM.

In this paper, we initiate a systematic study of the power of randomness for this problem. Our main contribution is perhaps to throw some light on where lies the difficulty. We first give an analysis of McGregor's algorithm using the traditional Primal-Dual framework (see Appendix A). All lower bounds for deterministic algorithms (both for MCM and MWM) employ *growing trees*. That is, the input graph is a tree at every stage. It is then natural to start our investigation for this class of inputs. For this class, we give a randomized algorithm (that uses two bits of randomness) that is  $\frac{28}{15}$  competitive. While this result is modest, already the analysis is considerably more involved than the traditional primal dual analysis. In the traditional primal dual analysis of the matching problem, the primal charge (every selected edge contributes one to the charge) is distributed (perhaps equally) to the two end-points. In the online case, this is usually done as the algorithm proceeds. Our assignment depends on the structure of the final tree, so this assignment happens at the end. Our charge distribution is *not* symmetric. It depends on the position of the edge in the tree (we make this clear in the analysis) as also the behavior of neighboring edges. The main technical lemma shows that the charge distribution will depend on a neighborhood of distance at most four. We also note that these algorithms are (restricted versions of) randomized greedy algorithms even in the offline setting. Obtaining an approximation ratio less than two for general graphs, even in the offline setting is a notorious problem. See [7, 2] for a glimpse of the difficulty.

The optimal maximal matching algorithm for MCM, and McGregor's [6] optimal deterministic algorithm for MWM are both local algorithms. The choice of whether a new edge should be accepted or rejected is based only on the weight of the new edge and the weight of the conflicting edges, if any, in the current matching.

It is natural to add randomness to such local algorithms, and to ask whether they do better than the known deterministic lower bounds. An obvious way to add randomness is to accept/reject the new edge with certain probability,

which is only dependent on the new edge and the conflicting edges in the current matching. The choice of adding a new edge is independent of the previous coin tosses used by the algorithm. We call such algorithms *randomized local algorithms*. We show that randomized local algorithms cannot do better than optimal deterministic algorithms. This indicates that randomized algorithms may have to use dependent coin tosses to get better approximation ratios. Indeed, the algorithm by Epstein et al. does this. So do our randomized algorithms.

The randomized algorithm of Epstein et al. [3] works as follows. For a parameter  $\theta$ , they round the weights of the edges to powers of  $\theta$  randomly, and then they update the matching using a deterministic algorithm. The weights get distorted by a factor  $\frac{\theta \ln \theta}{\theta - 1}$  in the rounding step, and the deterministic algorithm has a competitive ratio of  $2 + \frac{2}{\theta - 2}$  on  $\theta$ -structured graphs, i.e., graphs with edge weights being powers of  $\theta$ . The overall competitive ratio of the randomized algorithm is  $\frac{\theta \ln \theta}{\theta - 1} \cdot \left(2 + \frac{2}{\theta - 2}\right)$  which is minimized at  $\theta \approx 5.356$ . A natural approach to reducing this competitive ratio is to improve the approximation ratio for  $\theta$  structured graphs. However, we prove that the competitive ratio  $2 + \frac{2}{\theta - 2}$  is tight for  $\theta$ -structured graphs, as long as  $\theta \geq 4$ , for deterministic algorithms.

One (minor) contribution of this paper is a randomized algorithms for MCM on paths, that achieves a competitive ratio of  $\frac{4}{3}$ , with a matching lower bound.

The other (minor) contribution of this paper is to highlight model specific bounds. There is a difference in the models in which the lower and upper bounds have been proved and this may be one reason for the large gaps.

## 2 Barely Random Algorithms for MCM

In this section, we present barely random algorithms, that is, algorithms that use a constant number of random bits, for MCM on growing trees.

The ideal way to read the paper, for a reader of leisure, is to first read our analysis of McGregor's algorithm (presented in Appendix A), then the analysis of the algorithm for trees with maximum vertex degree three (presented in Appendix B.2) and then this section. The dual variable management which is the key contribution gets progressively more complicated. It is local in the first two cases. The Appendix B.3 also gives an example which shows why a non-local analysis is needed. Here are the well known Primal and Dual formulations of the matching problem. The primal formulation is known to be optimum for bipartite graphs. For general graphs, odd set constraints have to be added. But they are not needed in this paper.

Primal LP	Dual LP
$\max \sum_e x_e$	$\min \sum_v y_v$
$\forall v : \sum_{e \in e} x_e \leq 1$	$\forall e : y_u + y_v \geq 1$
$x_e \geq 0$	$y_v \geq 0$

### 2.1 Randomized Algorithm for MCM on Growing Trees

In this section, by using only two bits of randomness, we beat the deterministic lower bound of 2 for MCM on growing trees.

---

**Algorithm 1** Randomized Algorithm for Growing Trees

---

1. The algorithm maintains four matchings:  $M_1, M_2, M_3$ , and  $M_4$ .
  2. On receipt of an edge  $e$ , the processing happens in two phases.
    - (a) **The augment phase.** The new edge  $e$  is added to each  $M_i$  in which there are no edges adjacent to  $e$ .
    - (b) **The switching phase.** For  $i = 2, 3, 4$ , in order,  $e$  is added to  $M_i$  (if it was not added in the previous phase) and the conflicting edge is discarded, provided it decreases the quantity  $\sum_{i,j \in [4], i \neq j} |M_i \cap M_j|$ .
  3. Output matching  $M_i$  with probability  $\frac{1}{4}$ .
- 

We begin by assuming (we justify this below) that all edges that do not belong to any matching are leaf edges. This helps in simplifying the analysis. Suppose that there is an edge  $e$  which does not belong to any matching, but is not a leaf edge. By removing  $e$ , the tree is partitioned into two subtrees. The edge  $e$  is added to the tree in which it has 4 neighboring edges. (There must be such a subtree, see next para.) Each tree is analysed separately.

We will say that a vertex(/an edge) is *covered* by a matching  $M_i$  if there is an edge in  $M_i$  which is incident on(/adjacent to) the vertex(/edge). We also say that an edge is *covered* by a matching  $M_i$  if it belongs to  $M_i$ . We begin with the following observations.

- After an edge is revealed, its end points are covered by all 4 matchings.
- An edge  $e$  that does not belong to any matching has 4 edges incident on one of its end points such that each of these edges belong to a distinct matching. This holds when the edge is revealed, and does not change subsequently.

An edge is called *internal* if there are edges incident on both its end points. An edge is called *bad* if its end points are covered by only 3 matchings.

We begin by proving some properties about the algorithm. The key structural lemma that keeps “influences” of bad edges local is given below. The two assertions in the Lemma have to be proved together by induction.

**Lemma 2.1.**    1. *An internal edge is covered by at least four matchings (when counted with multiplicities). It is not necessary that these four edges be in distinct matchings.*

2. *If  $p, q$  and  $r$  are three consecutive vertices on a path, then bad edges cannot be incident on all 3 of these vertices, (as in figure 1).*

The proof of this lemma is in the Appendix B.4.

**Theorem 2.2.** *The randomized algorithm for finding MCM on growing trees is  $\frac{28}{15}$ -competitive.*

A local analysis like the one in Appendix B.2 will not work here. For a reason, see Appendix B.3. The analysis of this algorithm proceeds in two steps.

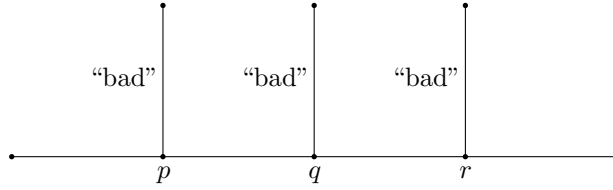


Figure 1: Forbidden Configuration

Once all edges have been seen, we impose a partial order on the vertices of the tree and then with the help of this partial order, we distribute the primal charge to the dual variables, and use the primal-dual framework to infer the competitive ratio. If every edge had four adjacent edges in some matching (counted with multiplicities) then the distribution of dual charge is easy. However we do have edges which have only three adjacent edges in matchings. We would like the edges in matchings to contribute more to the end-points of these edges. Then, the charge on the other end-point would be less and we need to balance this through other edges. Details follow.

**Ranks:** Consider a vertex  $v$ . Let  $v_1, \dots, v_k$  be the neighbors of  $v$ . For each  $i$ , let  $d_i$  denote the maximum distance from  $v$  to any leaf if there was no edge between  $v$  and  $v_i$ . The rank of  $v$  is defined as the minimum of all the  $d_i$ . Observe that the rank of  $v$  is one plus the second highest rank among the neighbors of  $v$ . Thus there can be at most one neighbor of vertex  $v$  which has rank at least the rank of  $v$ . All leaves have rank 0. Rank 1 vertices have at most one non-leaf neighbor.

**Lemma 2.3.** *There exists an assignment of the primal charge amongst the dual variables such that the dual constraint for each edge  $e \equiv (u, v)$  is satisfied at least  $\frac{15}{28}$  in expectation, i.e.  $\mathbb{E}[y_u + y_v] \geq \frac{15}{28}$ .*

*Proof.* Consider an edge  $e \equiv (u, v)$  where rank of  $u$  is  $i$  and rank of  $v$  is  $j$ . We will show that  $y_u + y_v \geq 2 + \epsilon$  for such an edge, when summed over all four matchings. The value of  $\epsilon$  is chosen later. The proof is by induction on the lexicographic order of  $\langle j, i \rangle$ ,  $j \geq i$ .

**Dual Variable Management:** Consider an edge  $e$  from a vertex of rank  $i$  to a vertex of rank  $j$ , such that  $i \leq j$ . This edge will distribute its primal weight between its end-points. The exact values are discussed in the proof of the claim below. In general, we look to transfer all of the primal charge to the higher ranked vertex. But this does not work and we need a finer strategy. This is detailed below.

- If  $e$  does not belong to any matching, then it does not contribute to the value of dual variables.
- If  $e$  belongs to a single matching then, depending on the situation, one of  $0$ ,  $\epsilon$  or  $2\epsilon$  of its primal charge will be assigned to the rank  $i$  vertex and rest will be assigned to the rank  $j$  vertex. The small constant  $\epsilon$  is determined later.
- If  $e$  belongs to two matchings, then at most  $3\epsilon$  of its primal charge will be assigned to the rank  $i$  vertex as required. The rest is assigned to the rank  $j$  vertex.

- If  $e$  belongs to three or four matchings, then its entire primal charge is assigned to the rank  $j$  vertex.

The analysis breaks up into six cases.

**Case 1.** Suppose  $e$  does not belong to any matching. Then it must be a leaf edge. Hence,  $i = 0$ . There must be 4 edges incident on  $v$  besides  $e$ , each belonging to a distinct matching. Of these 4, at least 3 say  $e_1$ ,  $e_2$ , and  $e_3$ , must be from lower ranked vertices to the rank  $j$  vertex  $v$ . The edges  $e_1$ ,  $e_2$ , and  $e_3$ , each assign a charge of  $1 - 2\epsilon$  to  $y_v$ . Therefore,  $y_u + y_v \geq 3 - 6\epsilon \geq 2 + \epsilon$ .

**Case 2.** Suppose  $e$  is a bad edge that belongs to a single matching. Since no internal edge can be a bad edge,  $i = 0$ . This implies (Lemma 2.1) that, there is an edge  $e_1$  from a rank  $j - 1$  vertex to  $v$ , which belongs to a single matching. Also, there is an edge  $e_2$ , from  $v$  to a higher ranked vertex, which also belongs to a single matching. The edge  $e$  assigns a charge of 1 to  $y_v$ . If  $e_1$  assigns a charge of 1 (or  $1 - \epsilon$ ) to  $y_v$ , then  $e_2$  assigns  $\epsilon$  (or  $2\epsilon$  respectively) to  $y_v$ . In either case,  $y_u + y_v = 2 + \epsilon$ . The key fact is that  $e_1$  could not have assigned  $2\epsilon$  to a lower ranked vertex. Since, then, by Lemma 2.1,  $e$  cannot be a bad edge.

**Case 3.** Suppose  $e$  is not a bad edge, and it belongs to a single matching. *Case 3(a).*  $i = 0$ . There are two sub cases.

- There is an edge  $e_1$  from some rank  $j - 1$  vertex to  $v$  which belongs to 2 matchings, or there are two other edges  $e_2$  and  $e_3$  from some lower ranked vertices to  $v$ , each belonging to separate matchings. The edge  $e$  assigns a charge of 1 to  $y_v$ . Either  $e_1$  assigns a charge of at least  $2 - 3\epsilon$  to  $y_v$ , or  $e_2$  and  $e_3$  assign a charge of at least  $1 - 2\epsilon$  each, to  $y_v$ . In either case,  $y_u + y_v \geq 3 - 4\epsilon \geq 2 + \epsilon$ .
- There is one edge  $e_1$ , from a rank  $j - 1$  vertex to  $v$ , which belongs to a single matching, and there is one edge  $e_2$ , from  $v$  to a higher ranked vertex, which belongs to 2 matchings. The edge  $e$  assigns a charge of 1 to  $y_v$ . If  $e_1$  assigns a charge of 1 (or  $1 - \epsilon$  or  $1 - 2\epsilon$ ) to  $y_v$ , then  $e_2$  assigns  $\epsilon$  (or  $2\epsilon$  or  $3\epsilon$  respectively) to  $y_v$ . In either case,  $y_u + y_v = 2 + \epsilon$ .

*Case 3(b).*  $i > 0$ . There are two sub cases.

- There are at least two edges  $e_1$  and  $e_2$  from lower ranked vertices to  $u$ , and one edge  $e_3$  from  $v$  to a higher ranked vertex. Each of these edges are in one matching only (not necessarily the same matching).
- There is one edge  $e_4$  from a vertex of lower rank to  $u$ , at least one edge  $e_5$  from a lower ranked vertex to  $v$ , and one edge  $e_6$  from  $v$  to a vertex of higher rank. All these edges belong to a single matching (not necessarily the same).

The edge  $e$  assigns a charge of 1 among  $y_u$  and  $y_v$ . If  $e_1$  and  $e_2$  assign a charge of at least  $1 - 2\epsilon$  each, to  $y_u$ , then  $y_u + y_v \geq 3 - 4\epsilon \geq 2 + \epsilon$ . Similarly, if  $e_4$  assigns a charge of at least  $1 - 2\epsilon$  to  $y_u$ , and  $e_5$  assigns a charge of at least  $1 - 2\epsilon$  to  $y_v$ , then  $y_u + y_v \geq 3 - 4\epsilon \geq 2 + \epsilon$ .

**Case 4.** Suppose  $e$  is a bad edge that belongs to two matchings. Then  $i = 0$ . This implies that there is an edge  $e_1$ , from  $v$  to a vertex of higher rank which belongs to a single matching. The edge  $e$  assigns a charge of 2 to  $y_v$ , and the edge  $e_1$  assigns a charge of  $\epsilon$  to  $y_v$ . Thus,  $y_u + y_v = 2 + \epsilon$ .

**Case 5.** Suppose  $e$  is not a bad edge and it belongs to two matchings. This means that either there is an edge  $e_1$  from a lower ranked vertex to  $u$ , which belongs to at least one matching, or there is an edge from some lower ranked vertex to  $v$  that belongs to at least one matching, or there is an edge from  $v$  to some higher ranked vertex which belongs to two matchings. The edge  $e$  assigns a charge of 2 among  $y_u$  and  $y_v$ . The neighboring edges assign a charge of  $\epsilon$  to  $y_u$  or  $y_v$  (depending on which vertex it is incident), to give  $y_u + y_v \geq 2 + \epsilon$ .

**Case 6.** Suppose,  $e$  belongs to 3 or 4 matchings, then trivially  $y_u + y_v \geq 2 + \epsilon$ . From the above conditions, the best value for the competitive ratio is obtained when  $\epsilon = \frac{1}{7}$ , yielding  $\mathbb{E}[y_u + y_v] \geq \frac{15}{28}$ .  $\square$

Lemma 2.3 implies that the competitive ratio of the algorithm is at most  $\frac{28}{15}$ .

### 3 Lower Bounds

#### 3.1 Lower Bound for MWM

In this section, we prove a lower bound on the competitive ratio of a natural class of randomized algorithms in the online preemptive model for MWM. The algorithms in this class, which we call *local* algorithms, have the property that their decision to accept or to reject a new edge is completely determined by the weights of the new edge and the conflicting edges in the matching maintained by the algorithm. Indeed, the optimal deterministic algorithm by McGregor [6] is a local algorithm. The notion of locality can be extended to randomized algorithms as well. In case of *randomized local algorithms*, the event that a new edge is accepted is independent of all such previous events, given the current matching maintained by the algorithm. Furthermore, the probability of this event is completely determined by the weight of the new edge and the conflicting edges in the matching maintained by the algorithm. Given that the optimal  $(3 + 2\sqrt{2})$ -competitive deterministic algorithm for MWM is a local algorithm, it is natural to ask whether randomized local algorithms can beat the deterministic lower bound of  $(3 + 2\sqrt{2})$  by Varadaraja [8]. We answer this question in the negative, and prove the following theorem.

**Theorem 3.1.** *No randomized local algorithm for the MWM problem can have a competitive ratio less than  $\alpha = 3 + 2\sqrt{2} \approx 5.828$ .*

Note that the randomized algorithm by Epstein et al. [3] does not fall in this category, since the decision of accepting or rejecting a new edge is also dependent on the outcome of the coins tossed at the beginning of the run of the algorithm. (For details, see Section 3 of [3].) In order to prove Theorem 3.1, we will crucially use the following lemma, which is a consequence of Section 4 of [8].

**Lemma 3.2.** *If there exists an infinite sequence  $(x_n)_{n \in \mathbb{N}}$  of positive real numbers such that for all  $n$ ,  $\beta x_n \geq \sum_{i=1}^{n+1} x_i + x_{n+1}$ , then  $\beta \geq 3 + 2\sqrt{2}$ .*

### 3.1.1 Characterization of local randomized algorithms

Suppose, for a contradiction, that there exists a randomized local algorithm  $\mathcal{A}$  with a competitive ratio  $\beta < \alpha = 3 + 2\sqrt{2}$ ,  $\beta \geq 1$ . Define the constant  $\gamma$  to be

$$\gamma = \frac{\beta(1 - \frac{1}{\alpha})}{(1 - \frac{\beta}{\alpha})} = \frac{\beta(\alpha - 1)}{\alpha - \beta} \geq 1 > \frac{1}{\alpha}$$

For  $i = 0, 1, 2$ , if  $w$  is the weight of a new edge and it has  $i$  conflicting edges, in the current matching, of weights  $w_1, \dots, w_i$ , then  $f_i(w_1, \dots, w_i, w)$  gives the probability of switching to the new edge. The behavior of  $\mathcal{A}$  is completely described by these three functions. We need the following key lemma to state our construction of the adversarial input.

The lemma states (informally) that given an edge of weight  $w_1$ , there exists weights  $x$  and  $y$ , close to each other such that if an edge of weight  $x$  (respective  $y$ ) is adjacent to an edge of weight  $w_1$ , the probability of switching is at most (respectively at least)  $\delta$ .

**Lemma 3.3.** *For every  $\delta \in (0, 1/\alpha)$ ,  $\epsilon > 0$ , and  $w_1$ , there exist  $x$  and  $y$  such that  $f_1(w_1, x) \geq \delta$ ,  $f_1(w_1, y) \leq \delta$ ,  $x - y \leq \epsilon$ , and  $w_1/\alpha \leq y \leq x \leq \gamma w_1$ .*

The proof of this lemma can be found in Appendix (section C).

### 3.1.2 The adversarial input

The adversarial input is parameterized by four parameters:  $\delta \in (0, 1/\alpha)$ ,  $\epsilon > 0$ ,  $m$ , and  $n$ , where  $m$  and  $n$  determine the graph and  $\delta$  and  $\epsilon$  determine the weights of its edges.

Define the infinite sequences  $(x_i)_{i \in \mathbb{N}}$  and  $(y_i)_{i \in \mathbb{N}}$ , as functions of  $\epsilon$  and  $\delta$ , as follows.  $x_1 = 1$ , and for all  $i$ , having defined  $x_i$ , let  $x_{i+1}$  and  $y_i$  be such that  $f_1(x_i, x_{i+1}) \geq \delta$ ,  $f_1(x_i, y_i) \leq \delta$ ,  $x_{i+1} - y_i \leq \epsilon$ , and  $x_i/\alpha \leq y_i \leq x_{i+1} \leq \gamma x_i$ . Lemma 3.3 ensures that such  $x_{i+1}$  and  $y_i$  exist. Furthermore, by induction on  $i$ , it is easy to see that for all  $i$ ,

$$1/\alpha^i \leq y_i \leq x_{i+1} \leq \gamma^i \quad (1)$$

These sequences will be the weights of the edges in the input graph.

Given  $m$  and  $n$ , the input graph contains several layers of vertices, namely  $A_1, A_2, \dots, A_{n+1}, A_{n+2}$  and  $B_1, B_2, \dots, B_{n+1}$ ; each layer containing  $m$  vertices. The vertices in the layer  $A_i$  are named  $a_1^i, a_2^i, \dots, a_m^i$ , and those in layer  $B_i$  are named analogously. We have a complete bipartite graph  $J_i$  between layer  $A_i$  and  $A_{i+1}$  and an edge between  $a_j^i$  and  $b_j^i$  for every  $i, j$  (that is, a matching  $M_i$  between  $A_i$  and  $B_i$ ).

For  $i = 1$  to  $n$ , the edges  $\{(a_j^i, a_{j'}^{i+1}) | 1 \leq j, j' \leq m\}$ , in the complete bipartite graph between  $A_i$  and  $A_{i+1}$ , have weight  $x_i$ , and the edges  $\{(a_j^i, b_j^i) | 1 \leq j \leq m\}$ , in the matching between  $A_i$  and  $B_i$ , have weight  $y_i$ . The edges in the complete graph  $J_{n+1}$  have weight  $x_n$ , and those in the matching  $M_{n+1}$  have weight  $y_n$ . Note that weights  $x_i$  and  $y_i$  depend on  $\epsilon$  and  $\delta$ , but are independent of  $m$  and  $n$ . Clearly, the weight of the maximum weight matching in this graph is bounded from below by the weight of the matching  $\bigcup_{i=1}^{n+1} M_i$ . Since  $y_i \geq x_{i+1} - \epsilon$ , we have

$$\text{OPT} \geq m \left( \sum_{i=1}^n y_i + y_n \right) \geq m \left( \sum_{i=2}^{n+1} x_i + x_{n+1} - (n+1)\epsilon \right) \quad (2)$$



The edges of the graph are revealed in  $n + 1$  phases. In the  $i^{\text{th}}$  phase, the edges in  $J_i \cup M_i$  are revealed as follows. The phase is divided into  $m$  sub phases. In the  $j^{\text{th}}$  sub phase of the  $i^{\text{th}}$  phase, edges incident on  $a_j^i$  are revealed, in the order  $(a_j^i, a_1^{i+1}), (a_j^i, a_2^{i+1}), \dots, (a_j^i, a_m^{i+1}), (a_j^i, b_j^i)$ .

### 3.1.3 Analysis of the lower bound

The overall idea of bounding the weight of the algorithm's matching is as follows. In each phase  $i$ , we will prove that as many as  $m - O(1)$  edges of  $J_i$  and only  $\delta m + O(1)$  edges of  $M_i$  are picked by the algorithm. Furthermore, in the  $i + 1^{\text{th}}$  phase, since  $m - O(1)$  edges from  $J_{i+1}$  are picked, all but  $O(1)$  edges of the edges picked from  $J_i$  are discarded. Thus, the algorithm ends up with  $\delta m + O(1)$  edges from each  $M_i$ , and  $O(1)$  edges from each  $J_i$ , except possibly  $J_n$  and  $J_{n+1}$ . The algorithm can end up with at most  $m$  edges from  $J_n \cup J_{n+1}$ , since the size of the maximum matching in  $J_n \cup J_{n+1}$  is  $m$ . Thus, the weight of the algorithm's matching is at most  $m x_n$  plus a quantity that can be neglected for large  $m$  and small  $\delta$ .

Let  $X_i$  (resp.  $Y_i$ ) be the set of edges of  $J_i$  (resp.  $M_i$ ) held by the algorithm at the end of input. Then we have,

**Lemma 3.4.** *For all  $i = 1$  to  $n$*

$$E[|Y_i|] \leq \delta m + \frac{1 - \delta}{\delta}$$

**Lemma 3.5.** *For all  $i = 1$  to  $n - 1$*

$$E[|X_i|] \leq \frac{1 - \delta}{\delta}$$

**Lemma 3.6.**

$$E[|Y_{n+1}|] \leq \delta m + \frac{1 - \delta}{\delta}$$

The proof of the above lemmas can be found in Appendix (section C).

We are now ready to prove Theorem 3.1. The expected weight of the matching held by  $\mathcal{A}$  is

$$E[\text{ALG}] \leq \sum_{i=1}^n y_i E[|Y_i|] + y_n E[|Y_{n+1}|] + \sum_{i=1}^{n-1} x_i E[|X_i|] + x_n E[|X_n \cup X_{n+1}|]$$

Using Lemmas 3.4, 3.6, 3.5, and the facts that  $y_i \leq x_{i+1}$  for all  $i$  and  $E[|X_n \cup X_{n+1}|] \leq m$  (since  $X_n \cup X_{n+1}$  is a matching in  $J_n \cup J_{n+1}$ ), we have

$$E[\text{ALG}] \leq \left( \delta m + \frac{1 - \delta}{\delta} \right) \left( \sum_{i=2}^{n+1} x_i + x_{n+1} \right) + \frac{1 - \delta}{\delta} \sum_{i=1}^{n-1} x_i + m x_n$$

Since the algorithm is  $\beta$ -competitive, for all  $n, m, \delta$  and  $\epsilon$  we must have  $E[\text{ALG}] \geq \text{OPT} / \beta$ . From the above and equation (2), we must have

$$\begin{aligned} \left( \delta m + \frac{1 - \delta}{\delta} \right) \left( \sum_{i=2}^{n+1} x_i + x_{n+1} \right) + \frac{1 - \delta}{\delta} \sum_{i=1}^{n-1} x_i + m x_n &\geq \frac{m}{\beta} \left( \sum_{i=2}^{n+1} x_i + x_{n+1} - (n + 1)\epsilon \right) \end{aligned}$$

Since the above holds for arbitrarily large  $m$ , ignoring the terms independent of  $m$  (recall that  $x_i$ 's are functions of  $\epsilon$  and  $\delta$  only), we have for all  $\delta$  and  $\epsilon$ ,

$$\delta \left( \sum_{i=2}^{n+1} x_i + x_{n+1} \right) + x_n \geq \frac{1}{\beta} \left( \sum_{i=2}^{n+1} x_i + x_{n+1} - (n+1)\epsilon \right)$$

that is,

$$x_n \geq \frac{1}{\beta} \left( \sum_{i=2}^{n+1} x_i + x_{n+1} - (n+1)\epsilon \right) - \delta \left( \sum_{i=2}^{n+1} x_i + x_{n+1} \right)$$

Taking limit inferior as  $\delta \rightarrow 0$  in the above inequality, and noting that limit inferior is super-additive we get for all  $\epsilon$ ,

$$\begin{aligned} \liminf_{\delta \rightarrow 0} x_n &\geq \frac{1}{\beta} \left( \sum_{i=2}^{n+1} \liminf_{\delta \rightarrow 0} x_i + \liminf_{\delta \rightarrow 0} x_{n+1} - (n+1)\epsilon \right) \\ &\quad - \limsup_{\delta \rightarrow 0} \delta \left( \sum_{i=2}^{n+1} x_i + x_{n+1} \right) \end{aligned}$$

Recall that  $x_i$ 's are functions of  $\epsilon$  and  $\delta$ , and that from equation (1),  $1/\alpha^i \leq x_{i+1} \leq \gamma^i$ , where the bounds are independent of  $\delta$ . Thus, all the limits in the above inequality exist. Moreover,  $\lim_{\delta \rightarrow 0} \delta \left( \sum_{i=2}^{n+1} x_i + x_{n+1} \right)$  exists and is 0, for all  $\epsilon$ . This implies  $\limsup_{\delta \rightarrow 0} \delta \left( \sum_{i=2}^{n+1} x_i + x_{n+1} \right) = 0$  and we get for all  $\epsilon$ ,

$$\liminf_{\delta \rightarrow 0} x_n \geq \frac{1}{\beta} \left( \sum_{i=2}^{n+1} \liminf_{\delta \rightarrow 0} x_i + \liminf_{\delta \rightarrow 0} x_{n+1} - (n+1)\epsilon \right)$$

Again, taking limit inferior as  $\epsilon \rightarrow 0$ , and using super-additivity,

$$\liminf_{\epsilon \rightarrow 0} \liminf_{\delta \rightarrow 0} x_n \geq \frac{1}{\beta} \left( \sum_{i=2}^{n+1} \liminf_{\epsilon \rightarrow 0} \liminf_{\delta \rightarrow 0} x_i + \liminf_{\epsilon \rightarrow 0} \liminf_{\delta \rightarrow 0} x_{n+1} \right)$$

Note that the above holds for all  $n$ . Finally, let  $\overline{x}_n = \liminf_{\epsilon \rightarrow 0} \liminf_{\delta \rightarrow 0} x_{n+1}$ . Then we have the infinite sequence  $(\overline{x}_n)_{n \in \mathbb{N}}$  such that for all  $n$ ,  $\beta \overline{x}_n \geq \sum_{i=1}^{n+1} \overline{x}_i + \overline{x}_{n+1}$ . Thus, by Lemma 3.2, we have  $\beta \geq 3 + 2\sqrt{2}$ .

### 3.2 Lower Bound for $\theta$ structured graphs

Recall that an edge weighted graph is said to be  $\theta$ -structured if the weights of the edges are powers of  $\theta$ . The following bound applies to any deterministic algorithm for MWM on  $\theta$ -structured graphs.

**Theorem 3.7.** *No deterministic algorithm can have a competitive ratio less than  $2 + \frac{2}{\theta-2}$  for MWM on  $\theta$ -structured graphs, for  $\theta \geq 4$ .*

The proof of the above theorem can be found in Appendix (section D).

## 4 Randomized Algorithm for Paths

When the input graph is restricted to be a collection of paths, then every new edge that arrives connects two (possibly empty) paths. Our algorithm consists of several cases, depending on the lengths of the two paths.

---

**Algorithm 2** Randomized Algorithm for Paths

---

```
1:  $M = \emptyset$ . { $M$  is the matching stored by the algorithm.}
2: for each new edge  $e$  do
3:   Let  $L_1 \geq L_2$  be the lengths of the two (possibly empty) paths  $P_1, P_2$  that
    $e$  connects.
4:   If  $L_1 > 0$  (resp.  $L_2 > 0$ ), let  $e_1$  (resp.  $e_2$ ) be the edge on  $P_1$  (resp.  $P_2$ )
   adjacent to  $e$ .
5:   if  $e$  is a disjoint edge  $\{L_1 = L_2 = 0\}$  then
6:      $M = M \cup \{e\}$ .
7:   else if  $e$  is revealed on a disjoint edge  $e_1$   $\{L_1 = 1, L_2 = 0, e_1 \in M\}$  then
8:     with probability  $\frac{1}{2}$ ,  $M = M \setminus \{e_1\} \cup \{e\}$ .
9:   else if  $e$  is revealed on an end point of a path of length  $> 1$   $\{L_1 > 1, L_2 = 0\}$ 
   then
10:    if  $e_1 \notin M$ ,  $M = M \cup \{e\}$ .
11:   else if  $e$  joins two disjoint edges  $\{L_1 = L_2 = 1, e_1, e_2 \in M\}$  then
12:    with probability  $\frac{1}{2}$ ,  $M = M \setminus \{e_1, e_2\} \cup \{e\}$ .
13:   else if  $e$  joins a path and a disjoint edge  $\{L_1 > 1, L_2 = 1, e_2 \in M\}$  then
14:    if  $e_1 \notin M$ ,  $M = M \setminus \{e_2\} \cup \{e\}$ .
15:   else if  $e$  joins two paths of length  $> 1$   $\{L_1 > 1, L_2 > 1\}$  then
16:    if  $e_1 \notin M$  and  $e_2 \notin M$ ,  $M = M \cup \{e\}$ .
17:   end if
18:   Output  $M$ .
19: end for
```

---

The following simple observations can be made by looking at the algorithm:

- All isolated edges belong to  $M$  with probability one.
- The end vertex of any path of *length*  $> 1$  is covered by  $M$  with probability  $\frac{1}{2}$ , and this is independent of the end vertex of any other path being covered.
- For a path of length 2, 3, or 4, each maximal matching is present in  $M$  with probability  $\frac{1}{2}$ .

**Theorem 4.1.** *The randomized algorithm for finding MCM on path graphs is  $\frac{4}{3}$ -competitive.*

The proof of above theorem can be found in Appendix (section E).

## References

- [1] Niv Buchbinder and Joseph Naor. The Design of Competitive Online Algorithms via a Primal-Dual Approach. *Foundations and Trends in Theoretical Computer Science*, 3(2-3):93–263, 2009.
- [2] T.-H. Hubert Chan, Fei Chen, Xiaowei Wu, and Zhichao Zhao. Ranking on Arbitrary Graphs: Rematch via Continuous LP with Monotone and Boundary Condition Constraints. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2014, Portland, Oregon, USA, January 5-7, 2014*, pages 1112–1122, 2014.

- [3] Leah Epstein, Asaf Levin, Danny Segev, and Oren Weimann. Improved Bounds for Online Preemptive Matching. In *30th International Symposium on Theoretical Aspects of Computer Science, STACS 2013, Kiel, Germany*, pages 389–399, 2013.
- [4] Joan Feigenbaum, Sampath Kannan, Andrew McGregor, Siddharth Suri, and Jian Zhang. On Graph Problems in a Semi-streaming Model. *Theor. Comput. Sci.*, 348(2):207–216, December 2005.
- [5] R. M. Karp, U. V. Vazirani, and V. V. Vazirani. An Optimal Algorithm for On-line Bipartite Matching. In *Proceedings of the Twenty-second Annual ACM Symposium on Theory of Computing, STOC '90*, pages 352–358, New York, NY, USA, 1990. ACM.
- [6] Andrew McGregor. Finding Graph Matchings in Data Streams. In *Proceedings of the 8th International Workshop on Approximation, Randomization and Combinatorial Optimization Problems, and Proceedings of the 9th International Conference on Randomization and Computation: Algorithms and Techniques, APPROX'05/RANDOM'05*, pages 170–181, Berlin, Heidelberg, 2005. Springer-Verlag.
- [7] Matthias Poloczek and Mario Szegedy. Randomized Greedy Algorithms for the Maximum Matching Problem with New Analysis. In *53rd Annual IEEE Symposium on Foundations of Computer Science, FOCS 2012, New Brunswick, NJ, USA*, pages 708–717, 2012.
- [8] Ashwinkumar Badanidiyuru Varadaraja. Buyback Problem - Approximate Matroid Intersection with Cancellation Costs. In *Automata, Languages and Programming - 38th International Colloquium, ICALP 2011, Zurich, Switzerland, Proceedings, Part I*, pages 379–390, 2011.
- [9] Andrew Chi-Chin Yao. Probabilistic computations: Toward a unified measure of complexity. In *18th Annual Symposium on Foundations of Computer Science, FOCS 1977*, pages 222–227, Oct 1977.

## Appendices

### A A Primal-Dual Analysis of a deterministic algorithm for MWM

In this section, we present a primal-dual analysis for the deterministic algorithm due to [6] for the maximum weight matching problem in the online preemptive model. The algorithm is as follows.

---

**Algorithm 3** Deterministic Algorithm for MWM

---

1. Fix a parameter  $\gamma$ .
  2. If the new edge  $e$  has weight greater than  $(1 + \gamma)$  times the weight of the edges currently adjacent to  $e$ , then include  $e$  and discard the adjacent edges.
- 

#### A.1 Analysis

**Lemma A.1.** [6] *The competitive ratio of this algorithm is  $(1 + \gamma)(2 + \frac{1}{\gamma})$ .*

We use the primal-dual technique to prove the same competitive ratio. This analysis technique is different from the one in [1]. In [1], the primal variables once set to a certain value are never changed whereas in our analysis the primal variables may change during the run of algorithm. The primal and dual LPs we use for the maximum weight matching problem are as follows.

Primal LP	Dual LP
$\begin{aligned} &\max \sum_e w_e x_e \\ &\forall v : \sum_{v \in e} x_e \leq 1 \\ &x_e \geq 0 \end{aligned}$	$\begin{aligned} &\min \sum_v y_v \\ &\forall e : y_u + y_v \geq w_e \\ &y_v \geq 0 \end{aligned}$

We maintain both primal and dual variables along with the run of the algorithm. On processing an edge, we maintain the following invariants.

- The dual LP is always feasible.
- For each edge  $e \equiv (u, v)$  in the current matching,  $y_u \geq (1 + \gamma)w(e)$  and  $y_v \geq (1 + \gamma)w(e)$ .
- The change in cost of the dual solution is at most  $(1 + \gamma)(2 + \frac{1}{\gamma})$  times the change in cost of primal solution.

These invariants imply that the competitive ratio of the algorithm is  $(1 + \gamma)(2 + \frac{1}{\gamma})$ .

We start with  $\vec{0}$  as the initial primal and dual solutions. Consider a round in which an edge  $e$  of weight  $w$  is given. Assume that all the above invariants hold before this edge is given. Whenever an edge  $e \equiv (u, v)$  is accepted by the algorithm, we assign values  $x_e = 1$  to the primal variable and  $y_u = \max(y_u, (1 + \gamma)w(e))$ ,  $y_v = \max(y_v, (1 + \gamma)w(e))$  to the dual variables of its end points.

Whenever an edge is rejected, we do not change the corresponding primal or dual variables. Whenever an edge  $e$  is evicted, we change its primal variable  $x_e = 0$ . The dual variables never decrease. Hence, if a dual constraint is feasible once, it remains so. We will now show that the invariants are always satisfied. These are three cases.

1. If the edge  $e \equiv (u, v)$  has no conflicting edges in the current matching, then it is accepted by the algorithm in current matching  $M$ . We assign  $x_e = 1, y_u = \max(y_u, (1 + \gamma)w(e))$  and  $y_v = \max(y_v, (1 + \gamma)w(e))$ . Hence,  $y_u \geq (1 + \gamma)w(e)$  and  $y_v \geq (1 + \gamma)w(e)$ . And hence, the dual constraint  $y_u + y_v \geq w(e)$  is feasible. The change in the dual cost is at most  $2(1 + \gamma)w(e)$ . The change in the primal cost is  $w(e)$ . So, the change in the dual cost is at most  $(1 + \gamma)(2 + \frac{1}{\gamma})$  times the change in the cost of the primal solution.
2. If the edge  $e \equiv (u, v)$  has conflicting edges  $X(M, e)$  and  $w(e) \leq (1 + \gamma)w(X(M, e))$ , then it is rejected by the algorithm. That happens when  $y_u + y_v \geq (1 + \gamma)w(X(M, e))$ , and the dual constraint for edge  $e$  is satisfied.
3. If the edge  $e \equiv (u, v)$  had conflicting edges  $X(M, e)$  and  $w(e) > (1 + \gamma)w(X(M, e))$ , then it is accepted by the algorithm in the current matching  $M$  and  $X(M, e)$  is/are evicted from  $M$ . We only need to show that the change in dual cost is at most  $(1 + \gamma)(2 + \frac{1}{\gamma})$  times the change in the primal cost. The change in primal cost is  $w(e) - w(X(M, e))$ . The change in dual cost is at most  $2(1 + \gamma)w(e) - (1 + \gamma)w(X(M, e))$ . Hence the ratio is at most

$$\begin{aligned}
& \frac{2(1 + \gamma)w(e) - (1 + \gamma)w(X(M, e))}{w(e) - w(X(M, e))} \\
&= 2(1 + \gamma) + \frac{(1 + \gamma)w(X(M, e))}{w(e) - w(X(M, e))} \\
&< 2(1 + \gamma) + \frac{w(e)}{w(e) - w(X(M, e))} \\
&\leq 2(1 + \gamma) + (1 + \frac{1}{\gamma}) \\
&= (1 + \gamma)(2 + \frac{1}{\gamma})
\end{aligned}$$

Here, the management of the dual variables was straight forward. The introduction of randomization complicates matters considerably, and we are only able to analyze the algorithm in the very restricted setting of paths and “growing trees”.

## B Barely Random Algorithms for MCM

### B.1 Randomized Algorithm for Paths

**Theorem B.1.** *The barely random algorithm for finding the MCM on paths is  $\frac{3}{2}$ -competitive, and no barely random algorithm can do better.*

We prove this theorem using the following lemma.

---

**Algorithm 4** Barely Random Algorithm for Paths

---

1. The algorithm maintains two matchings:  $M_1$  and  $M_2$ .
  2. On receipt of an edge  $e$ , the processing happens in two phases.
    - (a) The augment phase. Here, the new edge  $e$  is added to each  $M_i$  such that there is no edge in  $M_i$  sharing an end point with  $e$ .
    - (b) The switching phase. Edge  $e$  is added to  $M_2$  and the conflicting edge is discarded, provided it decreases the quantity  $|M_1 \cap M_2|$ .
  3. Output a matching  $M_i$  with probability  $\frac{1}{2}$ .
- 

**Lemma B.2.** *The dual constraint for each edge is satisfied at least  $\frac{2}{3}$ rd in expectation.*

$M_1$  and  $M_2$  are valid matchings and hence correspond to valid primal solutions. For each edge  $e \equiv (u, v)$  in some matching  $M_i$ , we distribute a charge of  $x_e = 1$  amongst dual variables  $y_u$  and  $y_v$  of its vertices. We prove that for each edge  $e$ ,  $y_u + y_v \geq \frac{4}{3}$ . Thus,  $\mathbb{E}[y_u + y_v] \geq \frac{2}{3}$ . Hence, this algorithm has a competitive ratio  $\frac{3}{2}$ . All the dual variables are initialized to 0. Suppose  $e \equiv (u, v) \in M_i$  for some  $i \in [2]$ . Then distribution of primal charge  $x_e$  amongst dual variables  $y_u$  and  $y_v$  is done as follows. If there is an edge incident on  $u$  which does not belong to any matching, and there is an edge incident on  $v$  which does belong to some matching, then  $e$  transfer a primal charge of  $\frac{2}{3}$  to  $y_u$  and rest is transferred to  $y_v$ . Else, the primal charge of  $e$  is transferred equally amongst  $y_u$  and  $y_v$ .

We look at three cases and prove that  $y_u + y_v \geq \frac{4}{3}$  for each edge  $e \equiv (u, v)$ .

1. The edge  $e$  is not present in any matching.
  - (a) If there are no edges on both its end points in the input graph, then this edge has to be covered by both  $M_1$  and  $M_2$ . So this case is not possible.
  - (b) If there is no edge on one end point (say  $u$ ) in the input graph, then  $e$  has to belong to belong to some matching. So, this case is not possible.
  - (c) If there are two edges incident on end points of  $e$  in the input graph, then each of them has to be covered by some matching. So,  $y_u + y_v \geq 2 \cdot \frac{2}{3} = \frac{4}{3}$ .
2. The edge  $e \equiv (u, v)$  is present in a single matching.
  - (a) If there are no edges on both its end points in the input graph, then this edge has to be covered by both  $M_1$  and  $M_2$ . So this case is not possible.
  - (b) If there is no edge on one end point (say  $u$ ) in the input graph, then edge on the other end point must be covered by the other matching. Otherwise, edge  $e$  would have been covered by both matchings. So,  $y_u + y_v \geq 1 + \frac{1}{3} = \frac{4}{3}$ .

- (c) If there two edges incident on end points of  $e$  in the input graph, then at least one of them has to be covered by other matching. Else, edge  $e$  would have been covered by both matchings. So,  $y_u + y_v \geq 1 + \frac{1}{3} = \frac{4}{3}$ .

3. The edge  $e \equiv (u, v)$  is present in both the matchings, then  $y_u + y_v = 2 \geq \frac{4}{3}$ .

This proves the above claim. The corollary of the above claim is that we have a  $\frac{3}{2}$ -competitive randomized algorithm for the MCM on paths.

*Proof.* (of the second part of Theorem B.1) Suppose  $U$  is the set of matchings used by a barely random algorithm  $\mathcal{A}$ . Following input is given to this algorithm. Reveal two edges  $x_1$  and  $y_1$  such that they share an end point. Let  $S$  be the set of matchings to which  $x_1$  is added, and  $\bar{S}$  be the set of matchings to which  $y_1$  is added. Here,  $U = S \cup \bar{S}$ . Now give two more edges  $x_2$  and  $y_2$  disjoint from the previous edges, such that  $x_2$  and  $y_2$  share an end point. Wlog,  $x_2$  will be added to set of matchings  $S$ , and  $y_2$  will be added to set of matchings  $\bar{S}$ . Give an edge between  $y_2$  and  $x_1$ . Continue the input similarly for  $i > 2$ . It can be seen that expected increase in the size of optimum matching is  $\frac{3}{2}$ , whereas increase in the size of matching held by the algorithm is 1. Thus, we get a lower bound  $\frac{3}{2}$  on the competitive ratio of any barely random algorithm.  $\square$

## B.2 Randomized Algorithm for Growing Trees with maximum degree 3

In this section, we give a barely random algorithm for growing trees, with maximum degree 3. We beat the lower bound of 2 for MCM on the performance of any deterministic algorithm, for this class of inputs. The edges are revealed in online fashion such the one new vertex is revealed per edge, (except for the first edge). Any vertex in the input graph has maximum degree 3.

---

### Algorithm 5 Randomized Algorithm for Growing Trees with $\Delta = 3$

---

1. The algorithm maintains 3 matchings  $M_1, M_2, M_3$ .
  2. On receipt of an edge  $e$ , the processing happens in two phases.
    - (a) The augment phase. Here, the new edge  $e$  is added to each  $M_i$  such that there is no edge in  $M_i$  sharing an end point with  $e$ .
    - (b) The switching phase. For  $i = 2, 3$ , in order,  $e$  is added to  $M_i$  and the conflicting edge is discarded, provided it decreases the quantity  $\sum_{i,j \in [3], i \neq j} |M_i \cap M_j|$ .
  3. Output a matching  $M_i$  with probability  $\frac{1}{3}$ .
- 

**Theorem B.3.** *The barely random algorithm for finding the MCM on growing trees with maximum degree 3 is  $\frac{12}{7}$ -competitive.*

We make following simple observations.

- There cannot be an edge which is not in any matching.



- Call an edge “bad” if its end points are covered by only two matchings. Indeed, an edge whose none of the end points are leaves, cannot be “bad”.
- An edge incident on a vertex of degree 3 cannot be “bad”, because there will be a distinct edge belonging to every matching.

We begin by proving a few simple lemmas regarding the algorithm.

**Lemma B.4.** *There cannot be “bad” edges incident on both vertices of an edge.*

*Proof.* Note that a “bad” edge is created when an edge is revealed on a leaf node of another edge which belongs to two or three matchings currently and finally belongs to only one matching.

Let  $e$  be an edge which currently belongs to three matchings, which means it is the first edge revealed. Now if an edge  $e_1$  is revealed on a vertex of  $e$ , then  $e_1$  would be added to one matching, and  $e$  would be removed from that matching, (in the switching phase of the algorithm). For  $e_1$  to be a bad edge,  $e$  should be switched out of one more matching. This can happen only if there are two more edges revealed on the other vertex of  $e$ . This means there cannot be “bad” edges on both sides of  $e$ .

Let  $e$  belongs to two matchings. Then  $e$  already has a neighboring edge  $e_2$  which belongs to some matching. When  $e_1$  is revealed on the leaf vertex of  $e$ , it will be added to one matching, in the augment phase. Now for  $e_1$  to be “bad”,  $e$  should switch out of some matching. This can only happen if there is one more edge  $e_3$  revealed on the common vertex of  $e$  and  $e_2$ . Again, the lemma holds.  $\square$

**Lemma B.5.** *If a vertex has three edges incident on it, then at most one of these edges can have a “bad” neighboring edge.*

*Proof.* Out of the three edges incident on a vertex, only one could have belonged to two matchings at any step during the run of algorithm. Hence, only that edge which belonged to two matchings at some stage during the run of algorithm can have a “bad” neighboring edge.  $\square$

*Proof.* (of Theorem B.3)  $M_1, M_2, M_3$  are valid matchings and hence correspond to valid primal solutions. For each edge  $e \equiv (u, v)$  in some matching  $M_i$ , we distribute a charge of  $x_e = 1$  amongst dual variables  $y_u$  and  $y_v$  of its end points. We prove that for each edge  $e$ ,  $y_u + y_v \geq \frac{7}{4}$ . Thus,  $\mathbb{E}[y_u + y_v] \geq \frac{7}{12}$ . Hence, this algorithm has a competitive ratio  $\frac{12}{7}$ . All the dual variables are initialized to 0. Suppose  $e \equiv (u, v) \in M_i$  for some  $i \in [3]$ . Then distribution of primal charge  $x_e$  amongst dual variables  $y_u$  and  $y_v$  is done as follows. If there is a “bad” edge incident on  $u$ , then edge  $e$  transfer  $\frac{3}{4}$  of its primal charge to  $y_u$  and rest of it to  $y_v$ . Else, edge  $e$  transfer its primal charge equally between  $y_u$  and  $y_v$ .

We look at three cases and then prove that  $y_u + y_v \geq \frac{7}{4}$  for each edge  $e \equiv (u, v)$ .

1. Edge  $e \equiv (u, v)$  is “bad”.  $e \in M_i$  for some  $i \in [3]$ .  $e$  will have some neighboring edge  $e_1$  such that  $e_1 \in M_j$  for  $j \in [3]$  and  $i \neq j$ . Let the common vertex between  $e$  and  $e_1$  be  $v$ . Then  $e_1$  will transfer  $\frac{3}{4}$  of its primal charge to  $y_v$ . Thus,  $y_u + y_v = \frac{7}{4}$ .
2. Edge  $e \equiv (u, v)$  is present in a single matching and not “bad”. This case has four sub cases.

- (a)  $e$  has one neighboring edge  $e_1$ . Then  $e_1$  should belong to two matchings.
- (b)  $e$  has two neighboring edges  $e_1$  and  $e_2$  both belonging to only one matching. If these are both on the same side of  $e$ , then at most one of them could have a “bad” neighboring edge (by lemma B.5). If these are on opposite sides of  $e$ , then none of them can have a “bad” neighboring edge.
- (c)  $e$  has three neighboring edges  $e_1$ ,  $e_2$ , and  $e_3$ , such that  $e_1$  and  $e_2$  are on one side of  $e$ , and  $e_3$  is on another side of  $e$ . At most one of  $e_1$  and  $e_2$  can have a “bad” neighboring edge (by lemma B.5).
- (d)  $e$  has four neighboring edges  $e_1$ ,  $e_2$ ,  $e_3$ , and  $e_4$ , such that  $e_1$  and  $e_2$  are on one side of  $e$ , and  $e_3$  and  $e_4$  are on another side of  $e$ .

We can see that in all the above sub cases,  $y_u + y_v \geq \frac{7}{4}$ .

- 3. Edge  $e \equiv (u, v)$  belongs to two or three matchings. Then,  $y_u + y_v \geq \frac{7}{4}$  trivially.

This proves that we have a  $\frac{12}{7}$ -competitive randomized algorithm for finding MCM on growing trees with maximum degree 3.  $\square$

### B.3 Example showing need of non-local analysis

Consider input graph as a 4-regular tree with large number of vertices, and an extra edge on every vertex other than the leaf vertices. Every edge other than the extra edges will belong some matching. For every edge that belongs to some matching, there will one edge on each of its end points which does not belong to any matching. If the rule for distributing primal charge among dual variables is similar to one described in section B.2, then for each edge belonging to some matching will transfer its primal charge equally amongst both its end points. For each edge which does not belong to any matching,  $y_u + y_v = 2$ , which will imply only a competitive ratio of 2. We wish to get a competitive ratio better than 2. So we need some other idea.

### B.4 Proof of Lemma 2.1

*Proof.* Consider an edge  $(u, v)$  revealed at  $u$ .

1. When revealed it is not put in any matching. This means that there are four covered edges incident on  $u$ . (Call an edge covered if it belongs to some matching.) This situation cannot change as more edges are revealed. Thus the edge will remain covered by four matchings, and can never become a bad edge.
2. When revealed it is put in one matching. This means that there are three matching edges on at least two covered edges incident on  $u$ . If there were three covered edges incident on  $u$  then they remain covered edges. So suppose otherwise. Then there are two covered edges of which one is in two matchings. Hence there will always be three matching edges covering  $u$ . If an edge is revealed at  $v$  then there will be four matching edges covering the given edge. The edge may become bad if  $v$  stays a leaf and if one of the matchings on the edge with two of them, switches.

3. When revealed it is put in two matchings. Then there are two matching edges at  $u$  and at least one covered edge. If there are two covered edges, they remain so. Of the two copies of the edge in matchings, one may switch to a new edge but will always remain adjacent to this edge. Hence there will always be three matching edges covering  $u$ . If an edge is revealed at  $v$  then there will be four matching edges covering the given edge. The edge may become bad if  $v$  stays a leaf and if one of the matchings on the edge with two of them, switches.
4. When revealed it is put in three matchings. Then there is one covered edge at  $u$ . If one more edge is now revealed on  $u$ , then we are back to case 3. If a new edge is revealed on  $v$ , it replaces  $(u, v)$  in one of the matchings. Now, even if more edges are revealed on either side of  $(u, v)$ , it continues to be covered by four matchings.
5. When revealed it is put in four matchings. If a new edge is revealed either on  $u$  or  $v$ , then this case reduces to case 2.

This completes the proof of the first part of lemma.

For the second part of lemma, consider a leaf edge present on each of the vertices  $p, q$ , and  $r$ . Suppose the leaf edge incident on  $q$  is bad. When this edge was revealed, there must have been some edge incident on  $q$ , either  $(p, q)$  or  $(q, r)$ , which belonged to two matchings. Wlog, assume  $(p, q)$  belonged to two matchings. Then for a matching to switch out this edge, there need to be three edges incident on  $p$ , and hence the leaf edge incident on  $p$  cannot be a bad edge.  $\square$

## C Proof of lemmas from section 3.1

**Lemma C.1.** *For every  $w > 0$ ,  $f_0(w) > 1/\alpha$ .*

*Proof.* If not, then a single edge of weight  $w$  results in algorithm's expected cost  $wf_0(w) \leq w/\alpha < w/\beta$ , whereas the optimum is  $w$ . This contradicts  $\beta$ -competitiveness.  $\square$

**Lemma C.2.** *For every  $w_1$  and  $w \leq w_1/\alpha$ ,  $f_1(w_1, w) = 0$ .*

*Proof.* If  $f_1(w_1, w) > 0$  for some  $w_1$  and  $w$  such that  $w \leq w_1/\alpha$ , then the adversary's input is a star, with a single edge of weight  $w_1$  followed by a large number  $n$  of edges of weight  $w$ . Regardless of whether the first edge of weight  $w_1$  is accepted or not, the algorithm holds an edge of weight  $w$ , with probability approaching 1 as  $n \rightarrow \infty$ , in the end. The optimum is  $w_1 \geq \alpha w > \beta w$ , thus, contradicting  $\beta$ -competitiveness.  $\square$

**Lemma C.3.** *For every  $w_1$ , and  $w \geq \gamma w_1$ ,  $f_1(w_1, w) \geq 1/\alpha$ .*

*Proof.* Suppose  $f_1(w_1, w) < 1/\alpha$  for some  $w_1$  and  $w$  such that  $w \geq \gamma w_1$ . The adversary's input is a star, with a large number  $n$  of edges of weight  $w_1$ , followed by a single edge of weight  $w$ . The algorithm must hold an edge of weight  $w_1$ , before the edge of weight  $w$  is given, with probability approaching 1 as  $n \rightarrow \infty$ . Therefore, in the end, the algorithm's cost is  $w$  with probability less than  $1/\alpha$ , and at most  $w_1$  otherwise. Thus, the expected weight of the edge held by the

algorithm is less than  $w/\alpha + w_1(1 - 1/\alpha)$ , whereas the adversary holds the edge of weight  $w$ . Since the algorithm is  $\beta$ -competitive and  $\beta < \alpha$ , we have

$$w < \beta \left[ \frac{1}{\alpha} \cdot w + \left( 1 - \frac{1}{\alpha} \right) w_1 \right] \Rightarrow \left( 1 - \frac{\beta}{\alpha} \right) w < w_1 \beta \left( 1 - \frac{1}{\alpha} \right) \Rightarrow w < \gamma w_1$$

This is a contradiction.  $\square$

*Proof of Lemma 3.3.* By Lemma C.2,  $f_1(w_1, w_1/\alpha) = 0$ , and by Lemma C.3,  $f_1(w_1, \gamma w_1) \geq 1/\alpha$ . Take a finite sequence of points, increasing from  $w_1/\alpha$  to  $\gamma w_1$ , such that the difference between any two consecutive points is at most  $\epsilon$ , and observe the value of  $f_1(w_1, z)$  at each such point  $z$ . Since  $f_1(w_1, w_1/\alpha) < \delta$  and  $f_1(w_1, \gamma w_1) > \delta$ , there must exist two consecutive points in the sequence, say  $y$  and  $x$ , such that  $f_1(w_1, y) \leq \delta$  and  $f_1(w_1, x) \geq \delta$ . Furthermore,  $x - y \leq \epsilon$  and  $w_1/\alpha \leq y \leq x \leq \gamma w_1$ , by construction.  $\square$

**Lemma C.4.** *For every  $i, j$ , the probability that  $a_j^i$  is not matched to any vertex in  $A_{i+1}$ , in the  $j^{\text{th}}$  sub phase of the  $i^{\text{th}}$  phase, just before the edge  $(a_j^i, b_j^i)$  is revealed, is at most  $(1 - \delta)^{m-j+1}$ .*

*Proof.* Consider the  $j^{\text{th}}$  sub phase of the  $i^{\text{th}}$  phase, in which, the edges  $(a_j^i, a_1^{i+1}), (a_j^i, a_2^{i+1}), \dots, (a_j^i, a_m^{i+1}), (a_j^i, b_j^i)$  are revealed. Before this sub phase, the number of unmatched vertices in  $A_{i+1}$  must be at least  $m - j + 1$ . Call this set  $A'$ . If  $a_j^i$  was matched at the end of phase  $i - 1$ , then the weight of edge incident on  $a_j^i$ , at the beginning of the current phase, is  $x_{i-1}$ . For each vertex  $a_{j'}^{i+1} \in A'$ , given that  $a_j^i$  did not get matched to any of  $a_1^{i+1}, \dots, a_{j'-1}^{i+1}$ , the probability that  $a_j^i$  gets matched to  $a_{j'}^{i+1}$  is  $f_1(x_{i-1}, x_i) \geq \delta$ . Thus, the probability of  $a_j^i$  not getting matched to any vertex in  $A' \subseteq A_{i+1}$ , in the current sub phase, is at most  $(1 - \delta)^{m-j+1}$ . Note that this argument applies even if  $a_j^i$  was not matched at the beginning of the current phase, due to Lemma C.1 and since  $\delta < 1/\alpha < f_0(x_i)$ .  $\square$

*Proof of Lemma 3.4.* First, observe that the sequence in which the edges are revealed ensures that no edge adjacent to any edge  $e \in M_i$  appears after  $e$ . Thus, if  $e$  is picked when it is revealed, it is never preempted, and is maintained till the end of input. Hence,  $Y_i$  is also the set of edges of  $M_i$  that were picked as soon as they were revealed.

When the edge  $(a_j^i, b_j^i)$  is given, the algorithm picks it with probability at most  $\delta$  (since  $f_1(x_i, y_i) \leq \delta$ ) if  $a_j^i$  was matched to some vertex in  $A_{i+1}$ . By Lemma C.4, the probability of  $a_j^i$  not being matched to any vertex in  $A_{i+1}$  is at most  $(1 - \delta)^{m-j+1}$ . Thus, the probability that the edge  $(a_j^i, b_j^i)$  appears in  $Y_i$  is at most  $\delta + (1 - \delta)^{m-j+1}$ . Hence,  $E[|Y_i|] \leq \delta m + \sum_{j=1}^m (1 - \delta)^{m-j+1} \leq \delta m + (1 - \delta)/\delta$ .  $\square$

*Proof of Lemma 3.5.* Consider the set  $A'$  of all vertices  $a_j^{i+1}$ , which remain matched to some vertex in  $A_i$  at the end of input. Then clearly,  $|A'| = |X_i|$ . Let us find the probability that a vertex  $a_j^{i+1}$  appears in  $A'$ . For this to happen, it is necessary that  $a_j^{i+1}$  not be matched to any vertex in  $A_{i+2}$ , in the  $j^{\text{th}}$  sub phase of the  $i + 1^{\text{st}}$  phase. By lemma C.4, this happens with probability at most  $(1 - \delta)^{m-j+1}$ . Thus,  $E[|X_i|] = \sum_{j=1}^m (1 - \delta)^{m-j+1} \leq (1 - \delta)/\delta$ .  $\square$

**Lemma C.5.** *For every  $j$ , the probability that  $a_j^{n+1}$  is not matched to any vertex in  $A_n \cup A_{n+2}$ , in the  $j^{\text{th}}$  sub phase of the  $n+1^{\text{st}}$  phase, just before the edge  $(a_j^{n+1}, b_j^{n+1})$  is revealed, is at most  $(1 - \delta)^{m-j+1}$ .*

*Proof.* This proof is analogous to the proof of Lemma C.4. If  $a_j^{n+1}$  was matched to some vertex in  $A_n$  at the end of the  $n^{\text{th}}$  phase, then it will continue to remain matched to some vertex in  $A_n \cup A_{n+2}$ , until the edge  $(a_j^{n+1}, b_j^{n+1})$  is revealed. Otherwise  $a_j^{n+1}$  will get matched to some vertex in  $A_{n+2}$  with probability at least  $1 - (1 - \delta)^{m-j+1}$ , and remain unmatched with probability at most  $(1 - \delta)^{m-j+1}$ .  $\square$

*Proof of Lemma 3.6.* This proof is analogous to the proof of Lemma 3.4. Again, the sequence in which the edges are revealed ensures that no edge adjacent to any edge  $e$  in any  $M_{n+1}$  appears after  $e$ . Thus, if  $e$  is picked when it is revealed, it is never preempted. Hence,  $Y_i$  is also the set of edges of  $M_i$  that were picked as soon as they were revealed.

When the edge  $(a_j^{n+1}, b_j^{n+1})$  is given, the algorithm picks it with probability at most  $\delta$  (since  $f_1(x_n, y_n) \leq \delta$ ) if  $a_j^{n+1}$  was matched to some vertex in  $A_n \cup A_{n+1}$ . Thus, the probability that the edge  $(a_j^i, b_j^i)$  appears in  $Y_{n+1}$  is at most  $\delta + (1 - \delta)^{m-j+1}$ . Hence,  $E[|Y_{n+1}|] \leq \delta m + \sum_{j=1}^m (1 - \delta)^{m-j+1} \leq \delta m + (1 - \delta)/\delta$ .  $\square$

## D Lower Bound for $\theta$ structured graphs

The overall idea of the adversarial strategy is as follows. The input graph is a tree whose edges are partitioned into  $n+1$  layers which are numbered 0 through  $n$  from bottom to top. Every edge in layer  $i$  has weight  $\theta^i$ . The edges are revealed bottom-up. The edges in layer  $i$  are given in such a manner that all the edges in layer  $i-1$  held by the algorithm will be preempted. This ensures that in the end, the algorithm's matching contains only one edge, whereas the adversary's matching contains  $2^{n-i}$  edges from layer  $i$ , for each  $i$ .

Let  $\mathcal{A}$  be any deterministic algorithm for maximum matching in the online preemptive model. The adversarial strategy uses a recursive function, which takes  $n \in \mathbb{N}$  as a parameter. For a given  $n$ , this recursive function, given by Algorithm 6, constructs a tree with  $n+1$  layers by giving weighted edges to the algorithm in an online manner, and returns the tree, the adversary's matching in the tree, and a vertex from the tree.

Let us prove a couple of properties about the behavior of the algorithm and the adversary, when the online input is generated by the call  $\text{MAKETREE}(n)$ .

**Lemma D.1.** *Suppose that the call  $\text{MAKETREE}(n)$  returns  $(T', M', v')$ . Then*

1.  $M'$  is a matching in  $T'$ .
2.  $M'$  does not cover the vertex  $v'$ .
3. The weight of  $M'$  is  $\sum_{i=0}^n \theta^i 2^{n-i} = (\theta^{n+1} - 2^{n+1})/(\theta - 2)$ .

*Proof.* By induction on  $n$ . For  $n=0$ , the claim is obvious from the description of  $\text{MAKETREE}$ . Assume that the claim holds for  $n-1$ , and consider the call  $\text{MAKETREE}(n)$ , which returns  $(T', M', v')$ . Then, by induction hypothesis, the two recursive calls must have returned  $(T_1, M_1, v_1)$  and  $(T_2, M_2, v_2)$  satisfying

---

**Algorithm 6** MAKE TREE( $n$ )

---

```
1: if  $n = 0$  then
2:   while true do
3:     Take fresh vertices  $v, v_1, v_2$ , and give the edges  $(v_1, v_2), (v, v_1)$  with
       weight 1.
4:      $T := \{(v_1, v_2), (v, v_1)\}$ .
5:     if algorithm picks  $(v_1, v_2)$  then
6:       return  $(T, \{(v, v_1)\}, v_2)$ 
7:     else if algorithm picks  $(v, v_1)$  then
8:       return  $(T, \{(v_1, v_2)\}, v)$ 
9:     else
10:      Discard  $T$  and retry.
11:    end if
12:  end while
13: else
14:  while true do
15:     $(T_1, M_1, v_1) := \text{MAKE TREE}(n - 1)$ 
16:     $(T_2, M_2, v_2) := \text{MAKE TREE}(n - 1)$ 
17:    Give the edge  $(v_1, v_2)$  with weight  $\theta^n$ .
18:    if algorithm picks  $(v_1, v_2)$  then
19:      Take a fresh vertex  $v$ , and give the edge  $(v, v_1)$  with weight  $\theta^n$ .
20:       $T := T_1 \cup T_2 \cup \{(v_1, v_2), (v, v_1)\}$ .
21:      if algorithm replaces  $(v_1, v_2)$  by  $(v, v_1)$  then
22:        return  $(T, M_1 \cup M_2 \cup \{(v_1, v_2)\}, v)$ .
23:      else
24:        return  $(T, M_1 \cup M_2 \cup \{(v, v_1)\}, v_2)$ 
25:      end if
26:    else
27:      {algorithm does not pick  $(v_1, v_2)$ }
28:      Discard the constructed tree and retry.
29:    end if
30:  end while
31: end if
```

---

the conditions of the lemma. Suppose the algorithm replaced  $(v_1, v_2)$  by  $(v, v_1)$  in its matching. Since  $v_1$  and  $v_2$  were respectively left uncovered by  $M_1$  and  $M_2$ ,  $M = M_1 \cup M_2 \cup \{(v_1, v_2)\}$  is a matching in  $T$ , and  $M'$  does not cover  $v' = v$ . The case when the algorithm did not replace  $(v_1, v_2)$  by  $(v, v_1)$  is analogous. In either case, the additional edge in  $M'$ , apart from edges in  $M_1$  and  $M_2$  has weight  $\theta^n$ , and  $M_1, M_2$  themselves have weight  $\sum_{i=0}^{n-1} \theta^i 2^{n-1-i}$ , by induction hypothesis. Thus, the weight of  $M'$  is  $\theta^n + 2 \sum_{i=0}^{n-1} \theta^i 2^{n-1-i} = \sum_{i=0}^n \theta^i 2^{n-i}$ .  $\square$

**Lemma D.2.** *When the call  $\text{MAKETREE}(n)$  returns  $(T, M, v)$ , the algorithm's matching contains exactly one edge from  $T$ . This edge is incident on  $v$  and has weight  $\theta^n$ .*

*Proof.* By induction on  $n$ . Again, the claim is obvious for  $n = 0$ . Assume that the claim holds for  $n - 1$ , and consider the call  $\text{MAKETREE}(n)$ , which returns  $(T', M', v')$ . At the end of the two recursive calls which return  $(T_1, M_1, v_1)$  and  $(T_2, M_2, v_2)$ . The algorithm will have exactly one edge  $e_1$  from  $T_1$  incident on  $v_1$ , and one edge  $e_2$  from  $T_2$  incident on  $v_2$ , by induction hypothesis. If the algorithm does not pick the next edge  $(v_1, v_2)$ , then the tree is discarded. If the algorithm picks that edge, then it must preempt  $e_1$  and  $e_2$ . Thereafter, if the algorithm replaces  $(v_1, v_2)$  by  $(v, v_1)$  in its matching, then  $v' = v$ . Otherwise, if the algorithm keeps  $(v_1, v_2)$ , then  $v' = v_2$ . In either case, the algorithm is left with exactly one edge, which is incident on  $v'$ , and which has weight  $\theta^n$ .  $\square$

The adversary's strategy is given by Algorithm 7, where  $n \geq 1$  is a parameter.

---

**Algorithm 7**  $\text{ADV}(n)$

---

```

1: while true do
2:    $(T_1, M_1, v_1) := \text{MAKETREE}(n - 1)$ 
3:    $(T_2, M_2, v_2) := \text{MAKETREE}(n - 1)$ 
4:   Give the edge  $(v_1, v_2)$  with weight  $\theta^n$ .
5:   if algorithm picks  $(v_1, v_2)$  then
6:     Take a fresh vertex  $v$ , and give the edge  $(v, v_1)$  with weight  $\theta^n$ .
7:     if algorithm replaces  $(v_1, v_2)$  by  $(v, v_1)$  then
8:       Take a fresh vertex  $v'$  and give the edge  $(v, v')$  with weight  $\theta^n$ .
9:        $T := T_1 \cup T_2 \cup \{(v_1, v_2), (v, v_1), (v, v')\}$ .
10:      return  $M_1 \cup M_2 \cup \{(v_1, v_2), (v, v')\}$ 
11:   else
12:     {algorithm still has  $(v_1, v_2)$ }
13:     Take a fresh vertex  $v'$  and give the edge  $(v_2, v')$  with weight  $\theta^n$ .
14:      $T := T_1 \cup T_2 \cup \{(v_1, v_2), (v, v_1), (v_2, v')\}$ .
15:     return  $M_1 \cup M_2 \cup \{(v, v_1), (v_2, v')\}$ 
16:   end if
17: else
18:   {algorithm does not pick  $(v_1, v_2)$ }
19:   Discard the constructed tree and retry.
20: end if
21: end while

```

---

**Lemma D.3.** *When a tree  $T$  is discarded in a call to  $\text{MAKETREE}(n)$  or  $\text{ADV}(n)$ ,  $\text{ALG}(T) \leq \left(2 + \frac{2}{\theta-2}\right) \cdot \text{ADV}(T)$ , where  $\text{ALG}(T)$  and  $\text{ADV}(T)$  are*

respectively the total weights of the edges of the algorithm's and the adversary's matchings, in  $T$ .

*Proof.* For  $n \geq 1$ , consider the two calls to MAKE TREE, which returned  $(T_1, M_1, v_1)$  and  $(T_2, M_2, v_2)$  before the edge  $(v_1, v_2)$  is revealed. By Lemma D.2, the algorithm had exactly one edge in each of  $T_1$  and  $T_2$ , and this edge had weight  $\theta^{n-1}$ . The tree was discarded because the algorithm did not pick the edge  $(v_1, v_2)$ . Thus,  $\text{ALG}(T) = 2\theta^{n-1}$ . On the other hand, the adversary picks the matching  $M_1 \cup M_2 \cup \{(v_1, v_2)\}$  which, by Lemma D.1, has weight  $\text{ADV}(T) = \theta^n + 2 \sum_{i=0}^{n-1} \theta^i 2^{n-1-i} = \sum_{i=0}^n \theta^i 2^{n-i} = (\theta^{n+1} - 2^{n+1})/(\theta - 2)$ . Thus,

$$\begin{aligned} \frac{\text{ADV}(T)}{\text{ALG}(T)} &= \frac{\theta^{n+1} - 2^{n+1}}{2\theta^{n-1}(\theta - 2)} = \frac{\theta}{2} \times \frac{1 - (\frac{2}{\theta})^{n+1}}{1 - \frac{2}{\theta}} \geq \frac{\theta}{2} \times \frac{1 - (\frac{2}{\theta})^2}{1 - \frac{2}{\theta}} \\ &= \frac{\theta}{2} \times \left(1 + \frac{2}{\theta}\right) = \frac{\theta}{2} + 1 \geq \left(2 + \frac{2}{\theta - 2}\right) \end{aligned}$$

The last inequality follows from the fact that  $\theta \geq 4$ . Finally, note that when the discard happens in a call to MAKE TREE(0),  $\text{ALG}(T) = 0$  and  $\text{ADV}(T) = 1$ .  $\square$

Now we are ready to prove Theorem 3.7.

*Proof of Theorem 3.7.* For  $n \geq 1$ , give the adversarial input by calling  $\text{ADV}(n)$ . If the call does not terminate, then an unbounded number of trees are discarded, and by Lemma D.3, a lower bound of  $\left(2 + \frac{2}{\theta - 2}\right)$  is forced on each discarded tree. If the call terminates, then suppose  $T$  is the final tree constructed. Let  $(T_1, M_1, v_1)$  and  $(T_2, M_2, v_2)$  be returned by the two calls to MAKE TREE( $n - 1$ ). By the description of  $\text{ADV}$  and Lemma D.2, it is clear that the algorithm holds only one edge of  $T$  in the end, and this edge has weight  $\theta^n = \text{ALG}(T)$ . On the other hand, the adversary's matching contains  $M_1$  and  $M_2$ , and two edges of weight  $\theta^n$ , where by Lemma D.1, the weight of  $M_1$  and  $M_2$  is  $(\theta^n - 2^n)/(\theta - 2)$  each. Thus,  $\text{ADV}(T) = 2\theta^n + 2(\theta^n - 2^n)/(\theta - 2)$ . Therefore,

$$\frac{\text{ADV}(T)}{\text{OPT}(T)} = 2 + \frac{2(\theta^n - 2^n)}{\theta^n(\theta - 2)} = 2 + \frac{2(1 - (\frac{2}{\theta})^n)}{\theta - 2}$$

This approaches  $\left(2 + \frac{2}{\theta - 2}\right)$  as  $n \rightarrow \infty$ . Furthermore, this lower bound is also forced on the trees discarded during the execution of  $\text{ADV}(n)$ . Thus, the algorithm can not have a competitive ratio less than  $\left(2 + \frac{2}{\theta - 2}\right)$ .  $\square$

## E Proof of Theorem 4.1

Theorem 4.1 can be proved using the following lemma.

**Lemma E.1.** *For any (maximal) path  $P$  of length  $n > 0$ ,*

- *if  $n$  is even then  $\mathbb{E}[|M \cap P|] \geq (3/4)(n/2) + 1/4 = p_0(n)$  (say).*
- *if  $n$  is odd then  $\mathbb{E}[|M \cap P|] \geq (3/4)(n/2) + 3/8 = p_1(n)$  (say).*



*Proof.* For  $n = 1$  and  $n = 2$ ,  $\mathbb{E}[|M \cap P|] = 1$ , and for  $n = 3$ ,  $\mathbb{E}[|M \cap P|] = \frac{3}{2}$ . Thus the lemma holds when  $n \leq 3$ . We will induct on the number of edges in the input. (Case  $n = 1$  covers the base case.) Suppose the lemma is true before the arrival of the new edge  $e$ . We prove that the lemma holds even after  $e$  has been processed. We may assume that the length  $n$  of the new path  $P$  resulting from addition of  $e$  is at least 4.

1. If  $n$  is even and  $L_2 = 0$ , (therefore  $L_1$  is odd, and  $L_1 \geq 3$ ),  $\Pr[e_1 \notin M] = \frac{1}{2}$ . Therefore,  $e$  is added to  $M$  with probability  $\frac{1}{2}$ .

$$\mathbb{E}[|M \cap P|] \geq p_1(n-1) + \frac{1}{2} \geq p_0(n)$$

2. If  $n$  is even,  $L_1 = n - 2$  and  $L_2 = 1$ .

$$\begin{aligned} \mathbb{E}[|M \cap P|] &\geq p_0(n-2) + 1 \\ &= \frac{3}{4} \left( \frac{n-2}{2} \right) + \frac{1}{4} + 1 \\ &\geq p_0(n) \end{aligned}$$

3. If  $n$  is even, and  $L_2 > 1$ , where  $n = L_1 + L_2 + 1$ ,  $L_1$  is even, and  $L_2$  is odd.  $\Pr[e_1 \notin M, e_2 \notin M] = \frac{1}{4}$

$$\begin{aligned} \mathbb{E}[|M \cap P|] &\geq p_0(L_1) + p_1(L_2) + \frac{1}{4} \\ &= \frac{3}{4} \left( \frac{L_1 + L_2 + 1}{2} \right) + \frac{1}{4} + \frac{3}{8} - \frac{3}{8} + \frac{1}{4} \\ &\geq p_0(n) \end{aligned}$$

4. If  $n$  is odd, and  $L_2 = 0$ , (therefore  $L_1$  is even, and  $L_1 \geq 3$ ),  $\Pr[e_1 \notin M] = \frac{1}{2}$ . Therefore,  $e$  is added to  $M$  with probability  $\frac{1}{2}$ .

$$\mathbb{E}[|M \cap P|] \geq p_0(n-1) + \frac{1}{2} = p_1(n)$$

5. If  $n$  is odd,  $L_1 = n - 2$  and  $L_2 = 1$ .

$$\begin{aligned} \mathbb{E}[|M \cap P|] &\geq p_1(n-2) + 1 \\ &= \frac{3}{4} \left( \frac{n-2}{2} \right) + \frac{3}{8} + 1 \\ &\geq p_1(n) \end{aligned}$$

6. If  $n$  is odd, and  $L_2 > 1$ , where  $n = L_1 + L_2 + 1$ ,  $L_1$  is even, and  $L_2$  is even.  $\Pr[e_1 \notin M, e_2 \notin M] = \frac{1}{4}$

$$\begin{aligned} \mathbb{E}[|M \cap P|] &\geq p_0(L_1) + p_0(L_2) + \frac{1}{4} \\ &= \frac{3}{4} \left( \frac{L_1 + L_2 + 1}{2} \right) + \frac{1}{4} + \frac{1}{4} - \frac{3}{8} + \frac{1}{4} \\ &= p_1(n) \end{aligned}$$

7. If  $n$  is odd, and  $L_2 > 1$ , where  $n = L_1 + L_2 + 1$ ,  $L_1$  is odd, and  $L_2$  is odd.  
 $\Pr[e_1 \notin M, e_2 \notin M] = \frac{1}{4}$

$$\begin{aligned}\mathbb{E}[|M \cap P|] &\geq p_1(L_1) + p_1(L_2) + \frac{1}{4} \\ &= \frac{3}{4} \left( \frac{L_1 + L_2 + 1}{2} \right) + \frac{3}{8} + \frac{3}{8} - \frac{3}{8} + \frac{1}{4} \\ &\geq p_1(n)\end{aligned}$$

This completes the induction and hence implies a  $\frac{4}{3}$ -competitive ratio for this algorithm.  $\square$